DOCUMENT RESUME

ED 055 974 SP 005 330

AUTHOR Howell, John J.
TITLE Performance Evaluation in Relation to Teacher Education and Teacher Certification.
PUB DATE Apr 71
NOTE 27p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Evaluation Criteria; *Evaluation Methods; *Performance Criteria; Problems; Teacher Certification; Teacher Education; *Teacher Evaluation; Teacher Role

ABSTRACT

Any particular teaching performance is limited by time and place and must be regarded as only a small sample of the large collection of activities called teaching. Evaluation is part of teaching, but it may also be a separate activity carried on by a third party. Evaluation methods differ according to the purpose, whether for guidance of preservice teachers, improvement of training programs, or certification. For both teacher education and teacher certification a major problem is that of defining all the kinds of jobs that teachers are expected to fill and adequately describing the social and institutional settings in which the jobs exist. No matter what the main reason for the evaluation of a particular teaching performance, the evaluator must take some account of what the teacher is trying to accomplish. If teaching performance is to be evaluated in detail, some theory of teaching is needed as a guide to how it should be broken down. Adequate justification for the theory must be available in some form. A particularly difficult problem in evaluating teaching performance is that of gathering sufficient accurate data. If teaching performance is to be judged on the basis of pupil learning, great care must be used to eliminate major sources of error in the assessment of learning. Among the many issues related to the problem of evaluating teacher performance are racial bias in teacher selection, merit rating, accountability, differentiated staffing, and community control of the schools. (RT)

ED055974

PERFORMANCE EVALUATION IN RELATION TO TEACHER EDUCATION AND TEACHER CERTIFICATION

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

by

John J. Howell

April, 1971

Division of Teacher Education
The City University of New York

SP005-330

PERFORMANCE EVALUATION IN RELATION TO TEACHER EDUCATION

AND TEACHER CERTIFICATION

Summary of the Salient Points

Because this paper examines aspects of the problem of evaluating teaching performance in some depth, a summary of some of the more salient points is here provided for convenience.

1. Any particular teaching performance is limited by time and place and other aspects of the immediate situation and must be regarded as only a small sample of the large collection of activities called teaching, which have direct or indirect effects on learning.

2. Evaluation, which involves defining or reviewing objectives, examination of experience, and making value judgments based jointly on objectives and information about experience, is part of teaching; but it may also be a separate activity carried on by a third party for purposes other than those belonging strictly to the teaching activity itself. Evaluation methods will differ according to its purposes, whether for guidance of teachers-in-preparation, improvement of training programs, or certification.

3. For both teacher education and teacher certification a major preliminary problem is that of defining all the kinds of jobs that teachers are expected to fill and adequately describing the social and institutional settings in which the jobs exist or will (or should) exist in the future.

4. No matter what the main reason for the evaluation of a particular teaching performance, the evaluator must take some account of what the teacher is trying to accomplish. So far as possible the evaluator must put this in the form of "behavioral objectives," if the teacher has not already done so, and where this is impossible he must deal with the problem of how to evaluate with less precise statements of purpose. Purposes may change and therefore the evaluator must check to see whether earlier statements of objectives still apply.

5. If teaching performance is to be evaluated in detail, some theory of teaching is needed as a guide as to how it should be broken down into its parts. The theory should show which parts are most important and why and how the parts combine to produce the result. Adequate justification for the theory must be available in some form. When the

evaluation is to be used for teacher certification, the reasoning behind it should support the conclusion that the entire procedure is sensible and either necessary or definitely beneficial.

6. A particularly difficult problem in evaluating teaching performance is that of gathering sufficient accurate data, pertinent to the purposes of the evaluation and free from the influence of factors over which the teacher could not be expected to have any control. If teaching performance is to be judged on the basis of pupil learning, great care must be used to eliminate major sources of error in the assessment of learning. Among the major sources of error are inadequate or poorly planned sampling and insufficient control over the conditions under which information is obtained. Instruments that have been developed and might be used in evaluation of teaching performance are discussed in another paper.

7. Among the many issues related to the problem of evaluating teaching performance are: racial bias in the selection of teachers, merit rating as applied to teachers, the accountability of persons and institutions for their educational effectiveness, the pros and cons of differential staffing of the schools, and the issue of community control of the schools.

I. Introduction

Although teacher education and teacher certification are obviously different and logically distinct, and serve purposes that are generally well understood, their mutual interdependence and the diversity of procedures that they employ often lead to confusion in discussions about them. Both are now in transition, under the combined impact of strong criticism and the availability of new major resources, and hence propositions about them that once could be taken for granted may no longer be true. It is no wonder, then, that the confusion proliferates. When the realities change, the language changes also: new terms and new meanings for older terms become popular and other words and phrases fade away. Semantic difficulties tend to become a serious problem for those who are attempting to deal with real needs in real situations on one hand and to communicate effectively

about them on the other. In these circumstances, it behooves us to be as explicit as possible about the meanings of the terms we use and about the assumptions we are making.

In the sections that follow we shall first consider the terminology currently being used in discussion of teacher education and certification, especially in discussions of "competency-based" programs. Next we shall attempt to identify the problems that are inherent in the evaluation of teaching performance. After that we shall consider, more briefly, some of the related issues in teacher education and teacher certification. In a separate paper the state of the art in instrumentation pertinent to performance evaluation will be discussed.

II. Some definitions of terms and exploration of concepts to which they refer

The indisputable fact of large-scale failure of schools to engender learning in their pupils at an acceptable rate has led critics of education to charge teachers with incompetence. Rather than attempting to deny it, educators generally have agreed that, whatever the reasons for it, teacher incompetence is very common and that drastic measures to remedy it are needed. Much effort, therefore, is being directed toward the development of teaching competence, and the terms "competency-based teacher training" and "competency-based teacher certification" are currently much in use. According to dictionary definitions, <u>competence</u> and <u>competency</u> have the same meaning, but in current educational usage there appears to be a difference. <u>Competence</u> is the more general term and refers to overall capability, or "answering all the requirements." It implies qualification by experience and training for a definite employment, and suggests general

personal fitness for it. Competency is used to refer to some particular element of competence. Contemporary writers often analyze teaching into specific tasks and use competency to refer to the capability, including perhaps a whole set of skills, necessary to the performance of a given task.

The concept of competence is, of course, not new and one of the semantic problems that arise comes from earlier usage and earlier ideas of how the fact of competence might be established. Historically, teaching competence has been inferred from such evidence as testimonials, letters of reference, transcripts of school records, degrees or diplomas, personality assessment, and cognitive tests and measurements. The current tendency is to rule out all such evidence and to depend solely on evidence derived from two sources: (1) evaluation of performance in real or realistic situations, and (2) inference from evidence of learning on the part of persons taught (Del Schalock, 1970). There is, however, a problem connected with the newer usage. While common sense would suggest that it must be easy to show that teacher competence has a major impact on pupil learning, the evidence of it from research is not abundant and, in particular, it lacks clarity and consistency across broad classes of teaching-learning situations (Flanders and Simon, 1969). The current concept of teacher competence depends more strongly on faith that its connections with pupil learning will be established by future research than on results already in hand from past research.

The connections between teaching competence and teaching performance also need to be considered. Teaching competence, as such, is not directly observable but is generally regarded as a more or less enduring personal characteristic; a specific teaching competency, too, is presumed to be

persistent and hence applicable to a whole series of similar situations within the limitations of its definition. A teaching performance, however, the observable manifestation of teaching competence, or competency, is bound by time and place and other general situational variables, which define its setting or context. Thus it provides an available test from which future performance in similar situations may be predicted, but no one performance or brief series of performances really takes the full measure of the competence or competency. That can only be determined in retrospect. Thus the term teaching performance, as we shall use it here, refers to a particular sample of the general class of activities called teaching without implying anything about the adequacy or representativeness of the sample. The schematic diagram in Fig. 1 illustrates these relationships. The solid arrows in the diagram signify causality, whereas the

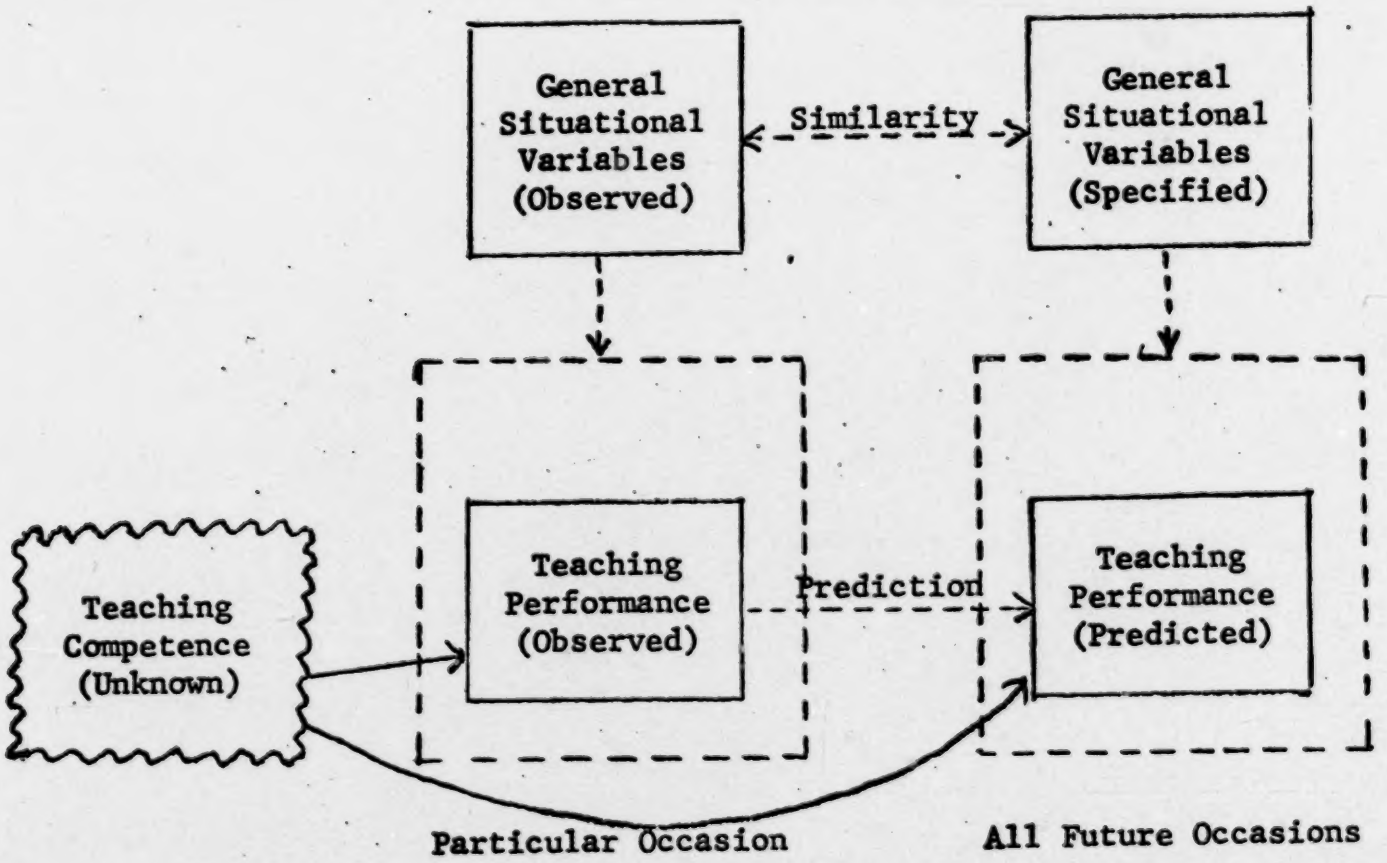

Fig. 1

broken arrows signify definition or inference. Temporal sequence from left to right is implied. The broken lines surrounding the boxes containing teaching performance represent the external context defined by what we are here calling general situational variables. The latter refer to the terms of reference for a teaching-learning situation as a whole, such as the institutional setting, the relatively fixed features of the physical surroundings, etc., as distinguished from "conditions of learning" subject to the teacher's immediate control. Generally the terms of reference, taken together over a period of time, are thought of by the practicing teaching as a particular teaching assignment, but for analytical purposes, they may be more narrowly drawn and used to define a teaching task --in which case, according to our previous definitions, "competence" would become "competency."

Any evaluation of teaching performance must begin with some concept of teaching, and evaluations based on radically different concepts may have little in common. Since teaching and related concept, learning, are the subjects of several highly developed theoretical systems, however, it would not be useful here to restrict either of them to a single arbitrary definition. Inasmuch as the conceptualization of teaching is one of the inherent problems of performance evaluation, it will be discussed more fully in a later section. Meanwhile, a few very general propositions about learning need to be considered.

Learning, like competence, is inferred from performance, not being directly observable itself. But since learning implies change--the development of some new behavior potential derived from experience--it can best be demonstrated by a comparison of two or more performances by the same learner. (Sometimes, however, it is inferred from a single performance

together with certain antecedent conditions and events, on the basis of an implicit or explicit assumption of zero potentiality for the performance in question at some earlier time.) The pupil performances in question imply, of course, contexts of their own, which are defined by what we will here call pupil situational variables. They provide the terms of reference for those observable behaviors from which learning is inferred. We note in passing that the term teaching does not enter into the above definition of learning, although we are here concerned primarily with learning that can be attributed to teaching. This rather gross account of the circumstances

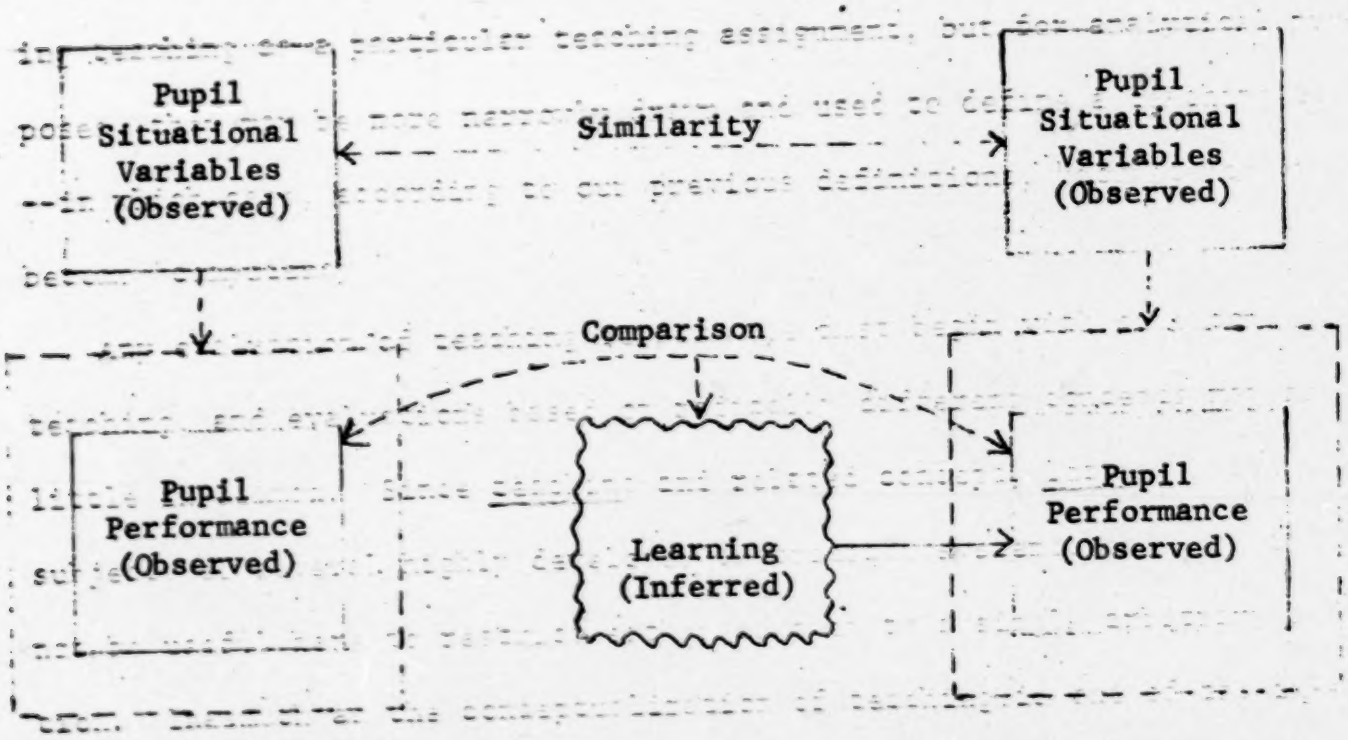

Fig. 2

that generally constitute the basis for inferring the occurrence of learning--it does not pretend to be a complete description--is illustrated by the schematic diagram in Fig. 2.

Let us now return to teaching and consider a general analysis, compatible, it is hoped, with several different theories of the teaching-learning

process, for the sake of clarity in the discussion of evaluation of teaching performance that will follow. Many concepts of teaching have emerged over the ages, but contemporary thought has been profoundly influenced by the behaviorist viewpoint in psychology with its insistent emphasis on description of operations and of observable conditions and events (Smith et al., 1969). In contemporary writing teaching appears to refer to a large collection of observable behaviors having direct or indirect effects on learning. (Sometimes it is defined to include behaviors merely intended to have effects on learning; it usually implies purposeful activity, which will be discussed further on.) The concept of teaching as "intervention," implying that learning may occur without it, is also current. From learning theories the concept of teaching has acquired the further connotations of being concerned with pupil learning behavior and of involving manipulation of conditions of learning. What is here called pupil learning behavior is distinct from the term pupil performance, which was used to denote the antecedent or subsequent series of pupil behaviors (e.g., in testing situations) from which the occurrence of learning could be detected. Because pupil learning behavior is overt and hence observable, it is also distinct from learning itself, which is not directly observable.

Usually teaching has been defined to include an evaluative function. Many contemporary writers see evaluation as part and parcel of the teaching-learning process (Burns et al., 1956). However, a separate role of "evaluator" is sometimes recognized and from this some confusion may arise. Because of the general concern about teaching and learning as matters pertinent to the public interest, and especially because of the pervasive feeling that they are not proceeding satisfactorily, many persons other than teachers have sought to determine for themselves what learning was

taking place. Thus the third-part concept of the evaluator emerged and a set of behaviors appropriate to the role of the outside evaluator, closely akin to the evaluation function in teaching, came to be recognized though not altogether consistently. The internal and external forms of evaluation need to be distinguished more clearly as to their uses and the circumstances of their occurrence, even though they are intrinsically not very different.

The term evaluation as applied to teaching is generally understood to include specification of objectives (purposes), examination of experience in terms of antecedent conditions, intervening actions, subsequent conditions, postulated cause-and-effect relationships, and value judgments based jointly on the information obtained and on the specified objectives. When viewed solely as a component of the total teaching-learning process, evaluation is seen as the concern of the persons involved. In some accounts it is treated primarily as the teacher's function, but others emphasize mutual teacher and pupil evaluation responsibilities and suggest exploration of the extent of correspondence between teacher and pupil purposes. The important point for the present discussion is that control over the information-getting processes necessary for evaluation rests with the teacher or with the teacher and the pupil(s). The conditions, events, and behaviors that constitute the information base may not be recorded--they may merely be observed and remembered--but explicit consideration of objectives, examination of experience, and formation of value judgments on the basis of both, untrammeled by any outside interference, are accomplished. The judgments that are the products of this kind of evaluation are useful to the teacher and pupil(s) in deciding their further strategies for teaching and learning, respectively.

The evaluation conducted by a third party, perhaps an outside evaluator, is similar in its processes and requires the same basic kinds of information, but typically employs more rigorous procedures and applies more stringent criteria of knowing. Toward this end the evaluator may attempt to extend his control, in particular, over the pupil situational variables that define the contexts for pupil performance from which learning may be inferred. In practice this may represent only a small incursion into the realm of teaching and may serve the teacher's purposes well, but the possibility of conflict between the teacher's and the evaluator's purposes is inherent in the social roles typically assigned to each. Frequently, when the evaluation is to be conducted in a school setting rather than in a laboratory, the solution with respect to control of information processes is to provide for an additional set of pupil performances, external to those included in teaching, and place them under the evaluator's jurisdiction. If that is done the information about pupil learning known to the evaluator is apt to be more detailed and precise than that ordinarily known to the teacher, but it may fail to include data pertinent to some of the teacher's objectives. Since the information sought by the evaluator depends on the purposes of the evaluation, these must be considered next.

We are not here concerned with all possible purposes of evaluation of teaching, but only with those that pertain to teacher preparation and teacher certification, which appear to be reducible to three:

1. For guidance of prospective teachers within the preparation program, i.e., formative evaluation with respect to individuals.
2. For guidance of teacher educators in program development or modification, i.e., formative evaluation with respect to programs, curricula, component activities, etc., in preparation for teaching.

3. For teacher certification.

While evaluation for the first two purposes can be viewed as internal to the teaching situation wherein the teacher educator is the "teacher" and the prospective teacher is the "pupil" they are clearly external to the prospective teacher's own teaching assignment in a laboratory or practice teaching situation. Here the teacher educator, or someone else, assumes the third-party role of evaluator and the evaluation methods more nearly resemble those of external evaluation than the more casual methods typically used by school teachers within the teaching process. Evaluation for teacher certification is always a third-party activity by reason of the accepted institutions, laws, and customs of our society. Logically it could hardly be otherwise, for there is nothing inherent in teaching that would regularly provide all the information needed for decision-making about certification.

Referring back to the earlier diagrams, it may be noted that, assuming third-party evaluation, the schema presented in Fig. 1 provides a rationale of teacher certification. With minor modification the schema in Fig. 2 provides a general model for the evaluation of any factor that affects learning.

III. Problems inherent in evaluation of teaching performance for purposes of teacher preparation and certification

Both teacher preparation and teacher certification require some overview of the social structure within which the future teacher will find employment and clear definitions of the requirements of the positions in which they will be placed. In truth, the word teacher is misleading because there are many kinds of teachers, some much more specialized than others, and good preparation for one kind of teaching position may be grossly inadequate for

another. Beyond this preliminary problem are a series of others having to do with purposes of teaching, with theory of the teaching-learning process, with experimental control of factors other than teaching which affect learning outcomes, and with practical difficulties in the execution of evaluation programs however well designed. These all have a bearing on the ultimate validity of the judgments and decisions derived from evaluation of teaching performance and applied to guidance of the prospective teacher, development and improvement of programs of preparation, or teacher certification. The overarching problem may be conceived as that of establishing validity in all phases of evaluation for each of these purposes.

1. Job definition

In the United States principal responsibility for teacher preparation and certification is vested in state educational agencies and hence, historically, the kinds of teaching assignments that have developed in practice have been confirmed and legitimized by such agencies (Stinnett, 1969). The lists of job titles provide important clues to the kinds of assignments a given person might be expected to undertake. (In New York State 57 different titles are listed, some with subclassifications.) A study of the job specifications for each position within some major jurisdiction, such as a state, would be a useful step toward realistic definition of the general situational variables for various kinds of teaching. For the practicing teacher these conditions are typically associated with acceptance of a job and include specification of the community, the place of teaching (ordinarily a particular area in a particular school), the nature of the group to be taught, the content areas for which the teacher is to be responsible, the time available for teaching, the salary schedule, etc. Because these

conditions are institutional, the individual teacher--especially the beginner--generally has to accept them as data, i.e., given conditions not readily subject to his control. While the planner for the future need not accept existing institutions as being immutable, he does need to have a good picture of what they are.

Although the types of teaching assignment for which prospective teachers are to be prepared and for which certification requirements are established may be clearly defined at a given time, they are not static in historical perspective (Andrews, 1970; Brubacker, 1947). A particular definition of job requirements may become obsolete, and hence the performance-based determination of competence, though valid at one time, may become inapplicable under new conditions. The problem this poses for the evaluator is that of deciding whether a given narrowly defined set of general situational variables is sufficiently stable to justify the effort and expense of basing a performance evaluation on them. The more broadly defined the job the less likely it is rapidly to become obsolete, but the subtle problems of gradual obsolescence affect even the jobs of generalists. The logical conclusions from these considerations in terms of teacher preparation and certification are: (a) preparation for a changing job must be a continual affair, and the performance criteria upon which preparation is based must themselves be subject to periodic review; (b) certification of specified kinds of teaching competence cannot in honesty promise the future relevance of that competence over a long term, as a lifetime teacher certification would appear to do.

While it is useful to know the present patterns of job specifications and the changes that are taking place in them, clarity about the desirable directions of social change is also needed. One aspect of the evaluator's

problem, therefore, is to conceptualize the entire set of teaching positions in some appropriate jurisdiction as they ought to be defined within some reasonable reality constraints. A basic issue is that of degree of specialization within the field of teaching. At the highest level of generality (lowest degree of specialization) one might conceive of an "all-purpose teacher" qualified to teach any or all subjects at any level to pupils of any characteristics and social origins in any kind of educational environment. Though such versatility of teaching competence may appear ridiculous and unlikely to occur in actual practice, this teacher generalist definition is often approximated in the breadth of teaching competence expected of the elementary school teacher and written into certification requirements. It seems likely that if flimsy evidence of competence were to be rejected, it would soon be apparent that the expectation of such a broad array of competencies in persons to be assigned to relatively low-ranking and low-paid positions was unrealistic. Either the teacher-generalist concept should be retained, in the institutional pattern envisaged for the future, and training and compensation should be adjusted upward in view of the diversity of demonstrated teaching competencies required of him, or else greater specialization among teachers of lower status should be considered. The low-paid generalist, expected to be all things to all men in all circumstances, can hardly fail to be accused of incompetence in many of the situations in which he may be placed, even though he may have a very respectable competence which is not being called upon. At present high degrees of teacher specialization are most typical at the secondary and advanced levels and especially in vocational fields.

2. Purposes of Teaching

After the teaching positions within which the competence or competencies will come into play have been adequately conceived in the social-

institutional sense a second problem arises--that of identifying and defining the purposes of teaching in the contexts in question. General teaching aims and purposes are strongly associated with the philosophic orientation of the persons or groups who define them and expressions of them in educational literature are not hard to find (Ammons, 1969). But their application to specific teaching assignments may not be explicit enough to guide an evaluation. And purposes are constantly in transition. This results in different and changing emphases among such outcomes of the teaching-learning process as motor skills, factual information, verbal communication skills, interests, esthetic appreciation and creativity, social conformity or nonconformity, critical thinking, and values or systems of belief.

In the definition of teaching purposes an issue arises with respect to "behavioral" statement. The translation of statements of teaching purpose into "behaviorally stated objectives" is essential to systematic and rigorous evaluation. Though evaluation in terms of more vaguely stated objectives is possible, it is likely to be impressionistic and subjective. It may depend on haphazard sources of information and thus be unverifiable. There are people, however, who object to the requirement that all teaching purposes must be put into the form of behavioral objectives on the grounds that it distorts those aims of teaching that refer more to the inward experience of the pupil than to his behavior. The effort to force all purposes of teaching into a behavioral mold may, according to some, be an impossible exercise which is not in the best interest of education in general (Andrews, 1969). The problem for the evaluator is to decide whether or not to include forms of evaluation that are impressionistic and of doubtful replicability as well as those that are rigorously objective.

The purposes of teaching may at times come into conflict with those of evaluation. Observation procedures essential to an evaluation may, for example, introduce pupil or teacher anxieties that are inimical to the development of desirable attitudes and behaviors. Creative expression in particular may be stifled by an evaluative frame of reference. Such conflicts may give rise to hostility between evaluator and teacher, especially when the roles are assigned to different people who have not fully understood and accepted each other's purposes. The challenge to the evaluator is to devise a way of accomplishing his own purposes without interfering with the accomplishment of those of the teacher.

A conceptual difficulty with regard to purposes of teaching is inherent in the question of whether to consider them as part of the external context or as part of the process of teaching. If one thinks of the purposes as being socially determined, in the main, and of the teacher as accepting the goal orientations of the surrounding community, acting as an agent of society in the education of its children, then one should presumably place purposes of teaching among the general situational variables that define the teacher's assignment. Some schools and some communities have a clearly identifiable philosophic climate in which expectations as to the emphases that teachers will place on various kinds of learning outcomes are explicit. But it is generally not true that purposes are merely given as an external reality to be accepted, for the teacher can influence them and so can the learner. If one concedes that they have an important effect on learning and also regards them as modifiable, one is led to place them among the conditions of learning. From this point of view such an activity as pupil-teacher planning (i.e., exploration, coordination, and harmonization of pupil and teacher purposes) is an integral part of teaching. Conceivably

different kinds of statements about teaching purposes might be included in both sets of variables with provision for systematic comparison so as to avoid contradictions. In any case the evaluator's problem, whether or not he is a third party, is to define objectives for use in evaluation which not only represent the expectations that others hold for the teacher and pupil but also sufficiently take into account the modifications of purpose that may grow out of the teaching-learning process itself. At the very least the evaluator needs to check from time to time the continuing validity of an original statement of purposes formulated, perhaps, some time in the past. There is also the possibility that retrospective evaluation may sometimes be needed based on purposes that have newly emerged.

3. Theory

The demands that evaluation makes upon teaching theory and learning theory, once the purposes of teaching have been defined, depend on the purposes of the evaluation. When the evaluation is intended to serve the needs of teachers-in-preparation for individual guidance or the needs of teacher educators for insights concerning program development, detailed analysis of teaching performance is important. There is a need to identify component parts and often it may be useful to focus on one part at a time. A theory provides the power to explain practice teaching, for instance, in terms of the reasons for the results obtained and hence to direct learner's and teacher's efforts to the particular aspects of performance most needing their attention.

When the teaching-learning process is treated analytically, the evaluator has the problem of justifying the theory on which the analysis is based. If established theory is invoked, presumably a body of research based on it may be drawn upon. Each component of teaching, however defined

and whatever its postulated relationship to other components, must have demonstrable validity in its own right if it is to be considered as a separate element of the teaching performance subjected to evaluation. Components that are not directly observable require evidence of construct validity--that is, they should be anchored to observable conditions or events both antecedent and subsequent to their own postulated occurrence. If, for instance, it is specified that teaching performance shall be judged in part on the basis of a task called "creation of a favorable climate for learning," then it must be shown (a) that desirable learning outcomes depend on favorable climate, and (b) that favorable climate depends on some set of teacher behaviors in some set of circumstances.

When the purpose of the evaluation is teacher certification only, the detailed analysis of teaching performance is not essential. The basic questions are whether, but not how, the intervention achieves its results, and whether a given performance, or set of performances, constitute valid evidence that similar results will regularly be obtained in the future by the individual candidate. While the evaluator may use a theory that subdivides teaching performance (e.g., by specifying "tasks," "strategies," "moves," etc.), he is not compelled to do so. If he does, however, he must call upon his theory for a rationale for combining them into a whole which can be evaluated and used in what is essentially a "go or no-go" decision. (Admittedly, the decision process can be fractionated into a series of decisions and earlier decisions in the series can be made tentative and can be used in individual guidance in a way that may temper the harshness of their impact. The point here is that the final decisions, and logically the certification process itself, rests on a summative evaluation and that therefore the validity of the formula for combining parts can

be brought into question and must be justified.)

One popular approach to evaluation for teacher certification is to establish a list of "minimum competencies" and insist that each is essential and therefore that certification must be conditioned upon demonstration of all by the candidate. This amounts to a requirement that all teachers in a given category must be alike with respect to these minima, though not necessarily with respect to levels of competency above the minimum nor with respect to competencies other than those included in the reference list. There can be no doubt that this is a possible procedure, but there may be some argument about its desirability. The argument in its favor would appear to rest on the assumption that if a teacher is incompetent in a certain area the pupils will not learn in that area, i.e., that they have no alternative to learning by the mediation of that particular teacher. It is conceivable, however, that children may learn "on their own" or by way of the casual (or planned) intervention of adults other than the teacher to whom they are assigned (Stephens, 1967). Perhaps they have less need for protection from specific areas of incompetence in an individual teacher than from general mediocrity, which certification based on a check list of minimum competencies would do nothing to prevent. This line of thinking might lead to the alternative approach of setting some minimum level for a combination of competencies, or a sort of index of general competence, without stipulation as to minimum levels of specific ingredients. The component competencies could be measured and reported without being used separately as criteria for certification, and thus school authorities in staffing could seek a balance in a whole group of teachers not required of each individual. The very talented person with one or two areas of glaring incompetence would not be barred, even from teaching positions in which

competence in those particular areas was ordinarily to be expected, but if employed despite the known weaknesses would presumably be assigned in such a way that the needs of pupils would be properly provided for. Whatever the certification procedures, the rationale on which they are based needs to be validated not merely as a possible and plausible one but as a necessary or definitely beneficial one. To be justifiable as public policy it must demonstrate that it tends to encourage persons with natural or acquired proficiencies for teaching to enter the profession, while persuading those lacking them and unlikely to acquire them to turn elsewhere.

4. Control of Extraneous Factors

Up to this point we have stressed the importance of locating teaching positions within the social-institutional context, of identifying teaching purposes and translating them into behavioral objectives so far as reasonably possible, and of applying valid theory to the analysis of teaching performance and to the processes of teacher preparation and certification. We now must go on to the problems of getting the detailed information the evaluator must have about specific instances of teaching performance in order to be able to make judgments or reach conclusions about them. Because of the fluidity typical of the teaching-learning process, it is a long step from merely knowing that it is happening to having systematic and pertinent information about it. The evaluator must resort to carefully devised procedures for observation and instruments for measurement of the variables defined by the theory he is using. The topic of instrumentation and the problems of validity associated with it will be discussed in a separate paper. But a related question will be considered now--that of the exclusion from the information compiled of contamination from factors other than teaching performance which might affect the observed out-

comes. For instance, if in the performance of a given task from which a particular teaching competency is to be inferred certain cues from the teacher would normally lead to a given type of pupil response but failed to do so because of undetected contrary influences, the interpretation encoded as information might be false.

In obtaining adequate information for the evaluation of teaching performance the evaluator has a twofold problem. First, with respect to the teaching performance itself he must identify and sufficiently describe all factors likely to have major effects on the learning in question, because potent factors, if unrecognized, may easily falsify the inference that learning resulted (or failed to result) from the conditions and events observed. And secondly, he must be equally circumspect about the influences upon the pupil performances from which his knowledge about the occurrence of learning is derived. In other words, extraneous influences on the teaching-learning process and extraneous influences on either or both of the testings from which data are obtained might contaminate the information and hence the evaluation. The known sources of possible contamination can often be dealt with in designing the evaluation procedures, and unknown ones can be countered by sampling teaching performance generously and averaging results over a number of occasions or over many learners. But this may be expensive. The sample size, the sampling procedures, control over pupil situational variables to assure comparable conditions for the pre- and post-learning performances, and recognition of interventions other than teaching--all are problems of the validity of the data, which are quite distinct from problems of the validity of the theoretical constructs or of the teaching purposes previously considered. The problems of data will be revisited in connection with our consideration of instruments for use in eval-

uation of teaching performance.

5. Execution

The practical difficulties in the execution of evaluation procedures need not be considered at length here, as many of them are only too well known and are discussed elsewhere (Weiner and Howell, 1967). Problems arising from conflicting purposes must certainly rank at least equal to those of limitations of resources. Since teaching performance is of vital concern to so many different people, whose needs and interests are different, the tendency to prejudge results is a very common source of practical difficulties.

IV. Related issues

Although many issues are related to those that are inherent in the evaluation of teaching performance, we shall consider only a few that are currently receiving considerable attention by the teaching profession and the general public. They are: racial bias in the selection of teachers, merit rating as applied to teachers, the accountability of persons and institutions for their educational effectiveness, the pros and cons of new practices in differential staffing of the schools, and the issue of community control of the schools.

1. Racial bias

The social changes of the past two decades, and especially the changes in the ethnic composition of the population of the large cities of the United States reflecting migration, have accentuated the problems of racial and ethnic imbalances in city school systems that previously existed. (Suburban and rural areas are, of course, not immune, but it is generally

conceded that the problem is presented most acutely in large cities.) In particular, the imbalance between the racial and ethnic composition of the teaching staff and that of the pupil population in many or most city schools brings to the fore the question of whether the procedures for the recruitment of candidates for teacher training and those for teacher certification (or licensure) are free from racial or ethnic discrimination. It can be argued--and is being argued--that existing examination programs, which purport to evaluate the candidate's suitability for the teaching profession on the basis of tests and interviews, effectively exclude members of minority groups whose potential capability as teachers can be demonstrated. This problem, like others previously considered, can be stated in terms of validity, for racial bias--like any other kind of bias--impairs the validity of the evaluation in which it occurs. While the old mistakes are clear enough, the new ones we may be making as we devise new performance-based evaluation procedures are not so readily apparent. Are the competencies themselves defined in such a way as to reflect values not shared by certain minorities in the population? Are the procedures by which they may be assessed free from discriminatory provisions? It might be argued that performance-based teacher certification itself would be a remedy for racial and ethnic bias in the recruitment and selection processes, but this will be true only to the extent that the new set of operations is itself untainted.

2. Merit rating

An issue of longer standing than that of racial bias, at least so far as nationwide emphasis is concerned, is the question of merit rating. One view is that a teacher's retention in employment, acquisition of tenure, and salary advancement should be made contingent upon effective performance

as determined by procedures specified in some merit system. The various plans set up for the determination of merit are, in fact, the progenitors of today's programs for performance evaluation, and some of the old arguments still apply though others are outmoded because of advances in technology. Opposition to merit rating in the past has often been able to challenge particular plans successfully, precisely on the grounds of the unsoundness of their evaluation processes. Methods of rating teacher competence were almost routinely found in well-controlled research to be lacking in predictive power against criteria of pupil learning. This objection will fail once fully valid systems of evaluating over-all teacher competence, i.e., adequate fulfillment of all the essential responsibilities of the position, become established. But other arguments will persist because they are based on other grounds, such as the fear that merit rating will create a climate of fear and suspicion and thus have an unwholesome effect on relationships among teachers.

3. Accountability

Somewhat related to the issue of merit rating is that of accountability for educational results. In one sense it is not an issue, for everyone seems to agree that accountability per se is a "good thing." The problems arise when one attempts to define who must be held accountable to whom, for what, and in what manner. Since this topic is currently receiving intensive consideration elsewhere, we will merely note in passing that as the methodology of evaluation of teaching performance develops, it is likely to be snatched up and put to use in more general programs of system-wide, area-wide, or state-wide evaluation now being considered under the impetus of the interest in accountability.

4. Differential staffing

An issue closely related to the problem of defining various teaching positions in terms of job specifications within social-institutional contexts, which was discussed earlier, is the now somewhat abated controversy about differentiated staffing. The concept of introducing "teacher aides" into the school staffing pattern was put forward with some fanfare by the Ford Foundation's Fund for the Advancement of Education in the early 1950's. Almost at once there was extensive opposition to the idea, based largely on some of the conditions, such as increased class size, that were associated with it in the early experiments. Since that time the use of various kinds of school aides, teacher aides, educational assistants, etc., often grouped under the generic term paraprofessionals, has increased rapidly and some of the controversy has subsided, though the issues have not necessarily been resolved. Since the newer patterns of differentiated staffing involve questions of preparation for and societal legitimization of the new positions and also the possibility of progression from paraprofessional to professional responsibilities (the "career ladder" concept), they cannot help affecting and being affected by developments in the area of evaluation of teaching performance.

5. Community control of schools

The final issue to be mentioned briefly here is that of community control of the schools. Sometimes this appears in the form of controversy over a plan of decentralization, as it has in New York City. Sometimes it relates mainly to questions of representation on boards of education, and nearly always it involves minority group interests and values. Its relevance

to evaluation of teaching performance is identical with its relevance to all aspects of evaluation: almost axiomatically any shift in the processes of social control and in the balance of representation within the structures where control is exercised entails corresponding shifts in purposes, which are the foundation upon which all evaluations are built. Furtherfore, changes in control almost inevitably affect the mechanisms of evaluation, its instrumentalities, procedures, and technology, because the latter are dependent to a considerable extent on the various sets of regulations and understandings through which general control is maintained.

References

Ammons, Margaret. Objectives and Outcomes, in Ebel, Robert L., Encyclopedia of Educational Research, Fourth Edition. New York: Macmillan, 1969, pp. 908-14.

Andrews, Theodore E. New Directions in Certification in Washington. (Paper submitted to: Improving State Leadership in Education Office, Denver, Colorado, September 1, 1970.) 75 pp.

Brubacher, John S. A History of the Problems of Education. Englewood Cliffs: Prentice Hall, 1947.

Burns, Constance M., et al. Yale-Fairfield Study of Elementary Teaching, Report for 1954-55, New Haven, 1956. 321 pp. (see Chapter V, pp. 106-139).

Del Schalock, H. The Focus of Performance-Based Certification: Knowledge, Teaching Behavior, or the Products That Derive from a Teacher's Behavior. (A paper prepared for the "Performance-Based Certification" Conference, The Florida State Department of Education, Miami Beach, Florida, May 19-22, 1970.) 10 pp.

Flanders, Ned A., and Simon, Anita. Teacher Effectiveness, in Ebel, Robert L., Encyclopedia of Educational Research, Fourth Edition. New York: Macmillan, 1969, pp. 1423-37.

Smith, B. Othanel, et al. Teachers for the Real World. Washington: American Association of Colleges for Teacher Education, 1969, 185 pp.

Stephens, J.M. The Process of Schooling, a Psychological Examination. New York: Holt, Rinehart and Winston, 1967, 168 pp.

Stinnett, T.M. Teacher Certification, in Ebel, Robert L., Encyclopedia of Educational Research, Fourth Edition. New York: Macmillan, 1969, pp. 1410-14.

Weiner, Max, and Howell, John. Problems of Evaluative Research in Large City Systems. (Presented at Northeastern Educational Research Association, November 16, 1970.) 7 pp.